Saumya Sinha | Curriculum Vitae

Boulder, Colorado

saumya.sinha@nrel.gov • https://saumyasinha.github.io/ • Google Scholar

I am a Postdoctoral Researcher at the National Laboratory of the Rockies (formerly named NREL) with a Ph.D. from the University of Colorado Boulder. My research focuses on leveraging advanced Machine Learning (ML) methods to address critical challenges in energy systems modeling and environmental science. My work has involved real-world applications such as physics-informed weather downscaling, generative models for scientific applications, extreme weather detection using satellite imagery, probabilistic solar energy forecasting, and predicting long-term sea level trends.

Education

- **University of Colorado, Boulder** — **Boulder, CO**
  PhD in Computer Science, with Prof. Claire Monteleoni — *2019 - 2023*
  Thesis title: Harnessing Deep Learning to Address Climate Change Impacts and Sustainability Challenges
- **University of Colorado, Boulder** — **Boulder, CO**
  Masters in Computer Science — *2017-2019*
- **Indian Institute of Technology, Kharagpur** — **Kharagpur, India**
  Integrated BS+MS in Mathematics and Computing — *2009–2014*

Employment

- **National Laboratory of the Rockies (formerly named NREL)** — **Golden, US**
  Postdoctoral researcher — *August 2023 - Present*
  Working in the Artificial Intelligence, Learning, and Intelligent Systems (ALIS) group in the Computational Science Center.
  Refer below to my publications in 2025 for my past work at the lab.
  Current projects include:
  - Developing a multimodal generative AI system for community-scale building energy modeling. This system leverages GPT-4o and the Segment Anything Model (SAM) to infer building attributes from satellite imagery and conversational inputs, generating simulation-ready data for energy load prediction.
  - Investigating how Large Language Model (LLM)-based assistants can combine domain knowledge with observational data to improve scientific reasoning, aiming to accelerate causal discovery for building interpretable probabilistic models.
- **AI for Good Research Lab, Microsoft** — **Redmond, US**
  Research Intern — *May 2022–Aug 2022*
  Self-supervised learning for geospatial imagery:
  - Made a contribution to *TorchGeo*, a PyTorch library for geospatial data.
  - Developed a self-supervised learning algorithm that uses a contrastive loss particularly designed to learn meaningful representations from satellite images and transfers well to semantic segmentation tasks. These representations were evaluated for the downstream task of land cover mapping.
- **American Express** — **Bangalore, India**
  Data Science Analyst, Big Data Capabilities and Algorithms division — *March 2016–May 2017*
  Text Mining/Modelling to build merchant database: Worked to improve the data quality of merchant (businesses using Amex services) attributes in a consumer-friendly database by integrating data from miscellaneous information sources.
- **Accenture** — **Bangalore, India**
  Data Science Analyst — *June 2014 - Feb 2016*
  Predictive Maintenance: Worked for an Oil and Natural Gas resources client, creating ML pipelines for failure detection of equipments in the plant.

Publications

Journals and Conferences

- **A Conditional Diffusion Model for Building Energy Modeling Workflows. (Under review, preprint link)**
  Saumya Sinha, Alexandre Cortiella, Rawad El Kontar, Andrew Glaws, Ryan King, Patrick Emami.
- **On the Effectiveness of Neural Operators at Zero-Shot Weather Downscaling. (link)**
  Saumya Sinha, Brandon Benton, Patrick Emami. *Environmental Data Science, 2025.*

- **SysCaps: Language Interfaces for Simulation Surrogates of Complex Systems.** **(link)**
  Patrick Emami, Zhaonan Li, **Saumya Sinha**, Truc Nguyen. *ICLR, 2025.*
- **Multi-decadal Sea Level Prediction using Neural Networks and Spectral Clustering on Climate Model Large Ensembles and Satellite Altimeter Data.** **(link)**
  Saumya Sinha, John Fasullo, Steve Nerem, Claire Monteleoni. *Artificial Intelligence for the Earth Systems, 2024.*
- **Week-ahead Solar Irradiance Forecasting with Deep Sequence Learning.** **(link)**
  Saumya Sinha, Bri-Mathias Hodge, Claire Monteleoni. *Environmental Data Science, 11th International Conference on Climate Informatics, 2022.*
- **Variational Autoencoder Anomaly-Detection of Avalanche Deposits in Satellite SAR Imagery.** **(link)**
  Saumya Sinha, Sophie Giffard-Roisin, Fatima Karbou, Michael Deschatres, Anna Karas, Nicolas Eckert, Cécile Coléou, Claire Monteleoni. *10th International Conference on Climate Informatics (ACM), 2020.*

Workshops

- **On the Effectiveness of Neural Operators at Zero-Shot Weather Downscaling.** **(link)**
  Saumya Sinha, Brandon Benton, Patrick Emami. *AAAI 2025 Workshop on AI to Accelerate Science and Engineering.*
- **SysCaps: Language Interfaces for Simulation Surrogates of Complex Systems.** **(link)**
  Patrick Emami, Zhaonan Li, **Saumya Sinha**, Truc Nguyen. *Neurips 2024 Workshop on Foundation Models for Science.*
- **Self-supervised Representation Learning of Geospatial Imagery for Semantic Segmentation.** **(Abstract link)**
  Saumya Sinha, Anthony Ortiz, Caleb Robinson, Rahul Dodhia, Juan Lavista Ferres. *AGU 2023.*
- **Sea level Projections with Machine Learning using Altimetry and Climate Model ensembles.** **(link)**
  Saumya Sinha, Steve Nerem, John Fasullo, Claire Monteleoni. *ICLR 2023 Workshop on Tackling Climate Change with ML.*
- **Sea-Level Projections via Spatiotemporal Deep Learning from Altimetry and CESM Large Ensembles.** **(Oral, Abstract link)**
  Saumya Sinha, Claire Monteleoni, John Fasullo, Steve Nerem. *AGU 2022: AI for Ocean and Climate Change session.*
- **Subseasonal Solar Power Forecasting via Deep Sequence Learning.** **(link)**
  Saumya Sinha, Bri-Mathias Hodge, Claire Monteleoni. *Neurips 2021 Workshop on Tackling Climate Change with ML.*
- **Detecting avalanche debris from SAR imaging: a comparison of convolutional neural networks and variational autoencoders.** **(Abstract link)**
  Sophie Giffard-Roisin, **Saumya Sinha**, Fatima Karbou, Michael Deschatres, Anna Karas, Nicolas Eckert, Cécile Coléou, and Claire Monteleoni. *EGU 2020.*
- **Detecting Avalanche Deposits using Variational Autoencoder on Sentinel-1 Satellite Imagery.** **(Spotlight Talk link)**
  Saumya Sinha, Sophie Giffard-Roisin, Fatima Karbou, Michael Deschatres, Anna Karas, Nicolas Eckert, Cécile Coléou, Claire Monteleoni. *Neurips 2019 Workshop on Tackling Climate Change with ML.*
- **Can Avalanche Deposits be Effectively Detected by Deep Learning on Sentinel-1 Satellite SAR Images?** **(link)**
  Saumya Sinha*, Sophie Giffard-Roisin*, Fatima Karbou, Michael Deschatres, Anna Karas, Nicolas Eckert, Cécile Coléou, Claire Monteleoni. *CI 2019 9th International Workshop on Climate Informatics.*

Professional service and Awards

- **Talks**:
  - Generative AI panelist in IEA Wind Task 43 Annual General Meeting 2025.
  - NREL AI Seminar 2025.
  - IBPSA-USA SimBuild 2024.
  - AI for Climate Science Bridge Program, AAAI 2023.
- **Awards at NREL and CU Boulder**:
  - Postdoctoral Researcher outstanding publication and innovation awards 2025.
  - Conference support fellowship 2022.
  - Departmental Summer Research fellowship for 2020-2021.
  - Annual Grad Students' Research Expo Award 2021.
  - Outstanding TA award for the year 2018-2019.
- **Reviewer (since 2019)**:
  - Journals: IEEE Transactions on Sustainable Energy, Journal of Renewable and Sustainable Energy, Transactions on Geoscience and Remote Sensing, Environmental Data Science.
  - Conferences: ICLR, ICML, Climate Informatics.
  - Program Committee member of ICLR & Neurips workshop on "Tackling Climate Change with Machine Learning".
- **Teaching**:
  - Guest Lecturer for ASEN5307 Engineering Data Analysis, CSCI 7000 Advanced Topics in Machine Learning, CSCI 4802/5802 Data Science Team.
  - Teaching Assistant for Machine Learning, Data Structures.